



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2000

Evaluating evaluations: does the Swiss practice live up to the ‘program evaluation standards’?

Widmer, Thomas

Abstract: In a meta-evaluation of fifteen Swiss evaluation studies, The Program Evaluation Standards have been applied as assessment criteria. The studies cover different policy fields (environmental, industrial, energy, social policy, and foreign policy). The analysis is based on written materials and qualitative interviews with both stakeholders and evaluators. The article assesses the relationship between The Program Evaluation Standards and current Swiss evaluation practice. It also highlights the factors that often prevent evaluators from undertaking thorough and provocative examinations in a small, consensus-oriented democracy without a long tradition in evaluation research and practice. Finally, the paper reflects on whether The Program Evaluation Standards are appropriate for a different sociocultural, political and policy context.

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-157861>

Book Section

Published Version

Originally published at:

Widmer, Thomas (2000). Evaluating evaluations: does the Swiss practice live up to the ‘program evaluation standards’? In: Russon, Craig. The program evaluation standards in international settings. Kalamazoo: Western Michigan University, Evaluation Center, 67-80.

Evaluating Evaluations: Does the Swiss Practice Live up to the ‘Program Evaluation Standards’?

Thomas Widmer
Institute of Political Science
University of Zurich

Abstract

In a meta-evaluation of fifteen Swiss evaluation studies, *The Program Evaluation Standards* have been applied as assessment criteria. The studies cover different policy fields (environmental, industrial, energy, social policy, and foreign policy). The analysis is based on written materials and qualitative interviews with both stakeholders and evaluators. The article assesses the relationship between *The Program Evaluation Standards* and current Swiss evaluation practice. It also highlights the factors that often prevent evaluators from undertaking thorough and provocative examinations in a small, consensus-oriented democracy without a long tradition in evaluation research and practice. Finally, the paper reflects on whether *The Program Evaluation Standards* are appropriate for a different sociocultural, political and policy context.

Paper presented at the 1. International Evaluation Conference, Vancouver, Canada, 1-5 November, 1995

1. The Swiss Political System and Evaluation

Evaluation, defined as a systematic inquiry assessing the outcomes of political programs, depends very heavily on its political context. The Swiss political system has some quite special features (compare Linder, 1994):

First, Switzerland has a three level government. Not only the Confederation, but also the 26 cantons (federal states) as well as the 3000 communities and municipalities have their own financial, legal and political resources. The power of government is therefore shared among the authorities on all three levels. “Most of the federal legislation is implemented by the cantons.” (Klöti, 1988: 194). The cantons have kept full autonomy in a limited number of areas (i.e. education, health, churches, taxes and political self-organization). Power sharing among the federal, cantonal, and municipal governments is very important. The organizing principle of this three-level system is subsidiarity. This means that the higher level has the power to intervene only in cases where the lower level is not capable of doing so successfully.

The second important characteristic of Switzerland is its small size (see Klöti, 1987: 2-4). Switzerland has approximately 6,700,000 inhabitants, which equals an average of 260,000 individuals per canton and 2,300 persons per municipality.

Third, Switzerland is characterized by a semi-direct democracy. Direct democracy has an important influence on the political system, namely on the decision rule: The dominant decision rule is consensus. On the federal level, the four political parties that have been represented in the federal government’s executive body in the same proportion since 1959 represent more than seventy per cent of the electorate.

The fourth aspect is the relative secrecy with which political decision-making takes place in Switzerland. According to the system of consensus, the principle of cooperativeness, called “Kollegialitätsprinzip” resp. “principe de collégialité”, prevents the individual members of the government from deviating from the official course of the body. In addition, the executive bodies of the federal government and almost all the cantons and communities do not apply the principle of non-restricted review of their decision-making process (“Öffentlichkeitsprinzip” resp. “principe de la publicité”).

In Switzerland, evaluation, and social science in general, has a relatively short history (see Bussmann, 1996: 307-308; Horber-Papazian and Thévoz, 1990). Until the mid-eighties, few in Switzerland talked about evaluation of the public sector. Since that time, new initiatives and changes have taken place. In 1987, the Federal Department of Justice and Police established a working group for legislative evaluations. This working group succeeded in contracting evaluation studies and further research projects on the use of systematic inquiry in Swiss legislation. In 1991, the group published its final report (AGEVAL, 1991). The report describes experiences in the evaluation process and contains some recommendations for practice and utilization of evaluation in Switzerland. In addition, the Swiss National Science Foundation started a national research program called *The Effectiveness of Public Programs* in 1989 (see

Bussmann 1996). Under the auspices of this program, a wide variety of evaluation studies was conducted to check the performance of different evaluation designs, methods, and techniques. The research program will be completed this year (see Bussmann, 1995; Bussmann et al. 1996). The meta-evaluative research project presented in the next chapter was a part of this program.

On the federal level, and to a lesser extent on the cantonal and local level, the use of evaluation has made some progress. First, the Federal Parliament has established a parliamentary administration control office whose duty it is to monitor administrative activities. This office acts only on the orders of the permanent commissions of the National Council and the Council of States. Since the new office was established, the parliament has used evaluation as a common instrument of administrative control. Second, the executive power now has its own office working on evaluation projects. The Federal Council's administration control office is in charge of monitoring administrative actions regarding effectiveness as well as efficiency.

Besides these two services, the government audit office has broadened the scope of its activities. The government audit office is increasingly involved in evaluative questions. In general, it is now more common to conduct evaluation studies in other branches of the central government. In the energy policy field, for example, many evaluations were carried out investigating the effects of the program called *Energy 2000*. This program promotes energy saving and the use of renewable energy sources. The evaluation studies analyze different program measures. The goals of the program were used as evaluative criteria. This is only one example of the widespread activities in evaluation that have taken place in the last few years.

In spite of the developments mentioned above, Switzerland still belongs to the countries lacking a strong evaluation community. Only a small group of social scientists is active in the field. There is little evaluation training in the social science curricula at the universities. There are no serious training programs at all for professional evaluators. A professional organization called Swiss Evaluation Society (SEVAL) has existed since 1995. Up to the present, the Society has been quite small and its activities have been rather limited. Switzerland still has to take many steps to become a country with a well-established evaluation tradition.

2. The Meta-evaluation: Research Questions and Design

The research project presented in this chapter was conducted from 1993 to 1995. It is called *Meta-evaluation—Criteria to Evaluate Evaluations* (see Widmer, 1996a&b). The central research question is divided into two parts:

- Does the Swiss practice live up to *The Program Evaluation Standards*?
- Are *The Program Evaluation Standards* appropriate for evaluating non-educational evaluations within the Swiss context?

Fifteen evaluation studies were selected from different policy fields in an effort to answer these questions. In a thorough analysis, each study was evaluated using an adapted version of *The*

Program Evaluation Standards, originally developed by the Joint Committee on Standards for Educational Evaluation (Joint Committee, 1981). Since the project started in 1993, it was not possible to use the recently published, second edition of *The Standards* (Joint Committee, 1994) as a reference in the empirical work. However, the 1994 edition was used to draw conclusions from the analyses. To facilitate the use of *The Standards*, they were adapted to Swiss conditions. This adaptation process included a translation into German and a slight adjustment to the non-educational and Swiss evaluation fields.

The resulting thirty criteria were used to evaluate each of the fifteen evaluations in an individual case study. The case studies are partially descriptive and partially evaluative. The descriptive part included information about the evaluand, the evaluative question, the evaluation design, the client, the evaluators, and a thick description of the evaluation process. The evaluative part contained a thorough assessment of the studies using the thirty evaluation criteria.

The alternative criteria proposed by Egon G. Guba and Yvonna S. Lincoln (1989: 236-243) were applied to those evaluation studies which did not follow a traditional scientific paradigm. These alternative criteria are more appropriate to evaluation studies that have an alternative paradigmatic orientation (e.g., naturalistic evaluation approaches). Both kinds of criteria were defined as idealistic criteria and were, therefore, hard to fulfill in reality: difficult circumstances confront every evaluation process (time and financial constraints, political pressure, poor data and so on). Therefore, in each case study, a short impressionistic (and subjective) judgment follows the assessment along the evaluative criteria. Furthermore, the case studies contain documentation which includes lists of the written sources consulted and of the interviewees.

The sample, consisting of the fifteen case studies, was drawn in two steps. A list of 110 evaluation studies conducted since 1990 constituted the universe for the selection of a sample of the first ten studies. The sampling procedure followed a set of theoretical dimensions. Among others, the following dimensions were considered: the governmental level, the institutional structure of the policy area, the financial, personnel, and time resources available to the evaluation, the methods applied, and the evaluator's professional and academic background. In this manner, the following ten evaluation studies were selected:

- Environmental policy and technological development, a study analyzing the effects of environmental regulation for heating systems on technological innovations in heating engineering (Balthasar and Knöpfel, 1994).
- *Innovation processes in enterprises*, evaluating the effects of public programs promoting economic development on innovation processes (Bierter and Binder, 1993).
- *Wage equality principle for homeworkers*, examining the effectiveness of a public program mandating equal wages for labor in factories and homework (Gerheuser and Schmid, 1993).

- *Housing program of the Federation*, an evaluation investigating the effects of a program to increase the owner-tenant ratio in the Swiss population (Schulz et al., 1993).
- *Community centers in Zurich*, analyzing the input-output efficiency of an organization that runs institutions open to the public promoting social life in the municipal districts of Zurich (Näf et al., 1993).
- *Clearing away the snow on the mountain passes in the canton of Uri*, a study searching for more efficient solutions to the problem of snow removal on the mountain passes in the canton of Uri (Sommer and Suter, 1993).
- *Fees for garbage disposal*, investigating whether the introduction of fees reduces the amount of garbage in Swiss municipalities (Haering et al., 1990).
- *The introduction of consumption-dependent charges for heating costs and the promotion of renewable energy sources*, examining the implementation and the effects of the two programs in three Swiss cantons (Linder et al., 1990).
- *Federal office for the environment, forests and landscapes*, an organizational evaluation assessing the organizational structure and the effects of the activities carried out by this federal office (Knoepfel et al., 1991).
- *The information campaign of the Federation concerning the European Economic Area-referendum*, a study dealing with the campaign of the federal government to promote the knowledge about the referendum on a Swiss EEA membership (Longchamp, 1993).

A further set of five studies, the second step in the sampling procedure, was chosen according to an additional research mandate. This mandate deals with the performance of evaluation studies under severe financial and time constraints (Widmer et al., 1996). The five studies were:

- *Investment aid in rural structural engineering in the canton of Zurich*, a study analyzing the effects of a cantonal program on structural change in agriculture (Schnyder, 1993).
- *Procedure of open bidding in the canton of Berne*, an investigation of the effects induced by the legislation on open bidding for contracts for public works and the possibilities for improvement (Suter and Walter, 1993).
- *Prevention of natural disasters in the canton of Wallis*, a project evaluating the measures taken to prevent damage by avalanches and landslides (Wuilloud, 1994).
- *Promotion of economic development in the canton of Neuchâtel*, an evaluation studying the diverse programs for promotion of regional economic development (Forster and Jeanrenaud, 1993).

- *Swiss design competition*, analyzing the economic and cultural impacts of a design competition called ‘Design Preis Schweiz’ (Thierstein, 1994).

The fifteen case studies, each dealing with one specific evaluation study, were based on written materials and qualitative interviews with both stakeholders and evaluators. The written materials consist of a wide variety of documents containing information about the evaluation. In addition to the evaluation report and its earlier draft(s), such materials include all correspondence related to the project, research designs, intermediate reports, the minutes, the evaluation contract (and if appropriate, contract revisions of it), reactions to the evaluation in the media or in other arenas. Interviewees were selected pragmatically. In all cases, the evaluator(s) and the client(s) were interviewed. Additional interviews were conducted with the most relevant stakeholders in the policy field under investigation.

The case studies were submitted to the evaluators, and if appropriate to the clients also, for their reaction to the arguments. On the basis of these statements, the case studies were revised. The evaluators were given the opportunity to explain any differing points of view in a formal statement, which was published along with the case studies (Widmer 1996a&b). The formulation of the case studies was, therefore, an interactive process with a strong negotiation component. In drawing the final conclusions, the case studies were compared with each other according to the evaluative criteria. The results of the meta-evaluation, presented in the next two chapters, answer the central research questions mentioned earlier.

3. Consequences for Evaluation Practice

The case studies revealed that the differences among the fifteen evaluation studies were considerable. One project was, in effect, not an evaluation, but an inventory of the program accomplishments (study number 13). In this case, the researchers refused to conduct a full case study and restricted themselves to a description of the project. Two other studies (number 5 and number 10) were labeled “evaluations” but were not evaluations at all in terms of the requirements imposed by the standards. They both should be called advisory studies rather than evaluations (compare Stanfield and Smith, 1984). The first conclusion follows from the above-mentioned observations: the term “evaluation” should be used carefully. The examples also demonstrate that evaluation in Switzerland is not yet a well-established profession.

Despite the fact that the meta-evaluation showed the evaluations were successful in terms of utility, feasibility and propriety, it also revealed several structural weaknesses. Table 1 shows the criteria with a positive or a negative assessment compared to the average.

Table 1: Criteria with positive or negative assessments (compare Joint Committee, 1981)

<i>Positive assessments</i>	<i>Negative assessments</i>
A.2. Evaluator Credibility	D.3. Described Purposes and Procedures
A.3. Information Scope and Selection	D.5. Valid Measurement
A.6. Report Dissemination	D.6. Reliable Measurement
A.7. Report Timeliness	D.10. Justified Conclusions
B.1. Practical Procedures	
C.1. Formal Obligation	
C.4. Public's Right to Know	
C.5. Rights of Human Subjects	

This survey highlights the discrepancy among the four criteria groups. On the left-hand side of the table, the reader will find the criteria of utility (A; four out of eight criteria), feasibility (B; one out of three criteria) and propriety (C; three out of eight criteria). On the right-hand side, in contrast, the reader will find the criteria dealing only with accuracy (D; four out of eleven). A comparable study on evaluation practice in the Netherlands came up with similar findings (see Algemene Rekenkamer, 1996).

The question, then, is how to eliminate the weaknesses detected in the accuracy domain without neglecting the strengths of practice in the other three areas. For this reason, it makes sense to take a closer look at the points attracting critique. Many of the evaluation studies analyzed did not report the procedures applied in the evaluation precisely enough. This is the reason why the evaluation process is not transparent enough to allow the reader's comprehension. This general weakness is more acute in the evaluations (or parts of them) employing qualitative data gathering or data analysis approaches. The explanation for the considerable qualitative-quantitative gap in procedural description has to do with the differing current levels of methodological development the two approaches have reached. The qualitative approaches, long neglected by mainstream evaluation (as in social science), have to make up for lost time. In addition, the inconsistencies between the requirements of a practical and a scientific report are apparent. The evaluators, criticized in this respect, often pointed out this dilemma. Evaluation has the obligation to combine the two requirements. Otherwise, evaluation risks losing its practical relevance or will not be considered for the serious peer-review that will ensure high-quality systematic inquiry. The two sides have to be considered in a balanced manner, without neglecting one or the other.

In addition to the developments in the Swiss evaluation domain described at the beginning of this paper, other steps should be taken toward professionalization. The authors suggest implementing plans for a strengthening of the Swiss Evaluation Society (SEVAL). Taking into account Switzerland's small size, it would perhaps be useful to consider stronger cooperation with the newly founded European Evaluation Society (EES). With the help of a strong professional organization it would be possible to establish an educational program dedicated to evaluation. Another tool in improving the professional self-image of evaluators in Switzerland would be the formulation of guiding principles (compare Shadish et al., 1995) or standards (see chapter 4). In the authors' opinion, it would be useful to conduct further meta-evaluations as a suitable instrument in quality assessment as well as in quality management.

4. Some Remarks about the Relationship between the Standards and Swiss Practice

The implementation of the translated *Program Evaluation Standards* was very fruitful. In general, there appears to be no reason to deny the applicability of the criteria to the Swiss situation. The complete coverage of all relevant topics is especially useful. The criteria set proved to be very accurate. In spite of this rather positive assessment, there are also some points of critique to mention:

1. It is awkward to use such a high number (30) of criteria. The authors believe it is important to try to reduce the number of criteria.
2. The set of criteria does not fulfill the requirements usually imposed on a social science typology. The criteria have many links among themselves; therefore, they are not independent of each other. In addition, many aspects are covered by more than one criterion. Also, not all the criteria are located on the same analytical level. This could be the reason why "chain judgment" occurs. Furthermore, the postulates formulated in the criteria conflict with each other. This leads to severe problems. As a consequence, it is nearly impossible to fulfill all the criteria if they are strictly applied. This is not a satisfactory situation for the evaluators responsible for the evaluation under investigation. But it represents the evaluator's dilemma very well.

The authors investigated the relationship among the thirty criteria in the empirical case studies with a network analysis. Looking at the newly-published revised edition of *The Program Evaluation Standards* (Joint Committee, 1994) in the light of this empirical results, the following suggestions are made regarding the criteria:

- Criterion U.4. "Valuational Interpretation" should be deleted from the list because of its strong interdependence with other criteria (especially to A.3. "Described Purposes and Procedures", A.10. "Justified Conclusions", and U.5. "Report Clarity") and because of the high number of criteria.
- The authors fully agree with the Joint Committee's combining U.7. "Report Timeliness" and U.6. "Report Dissemination" into one standard.

- Contrary to the Joint Committee's decision, the authors would like to keep to the older version of criterion F.3. "Cost Effectiveness." The new formulation is less clear than the older one and we think the new version tends to overemphasize the justification of the resources expended. Compared to the revised version, more attention should be paid to the provision of valuable information.
- The revision of standard P.4. "Public's Right to Know" to the new entry "Service Orientation" leads to problems in application in a non-organizational context. The authors suggest keeping the older version.
- In the Swiss situation, the change of P.7. "Balanced Reporting" to "Complete and Fair Assessment" seems to be of little use. While the broadening of the scope of the standard from "presentation" to "examination and recording" makes sense, the authors agree with Michael Q. Patton's critique on the impossibility of totally achieving completeness (Patton, 1994: 197).
- The authors fully approve of the modifications proposed by the Joint Committee to criteria A.1. "Object Identification". The title of the new version comes closer to describing the content of the standard and the formulation of the standard has gained in clarity.
- The two criteria dealing with validity and reliability, as the authors empirical research and methodological reflections have shown, are strongly interdependent. It would be a worthwhile improvement to combine the two into one standard. In this way, it would be possible to give a far-reaching statement on the quality of measurement.
- The new edition of the standards has an entirely new entry called "Metaevaluation". The authors judge this to be a sound addendum and disagree with the opinion expressed by Joe Hansen (1994: 559-560). First of all, Gene V. Glass calls a substantive combination of different studies a *meta-analysis* and not a *meta-evaluation* (see for example Smith/Glass, 1981). Secondly, there is a difference between an evaluation and an audit (compare Schwandt and Halpern, 1988). The third point is that metaevaluation's semantic meaning is an evaluation of an evaluation (compare the analogous formation of concepts as meta-communication, meta-theory and so on). Fourth, the term metaevaluation has often been used in the way proposed by the Joint Committee (see for example Scriven, 1969; Stufflebeam, 1978; Smith, 1981; Schwandt and Halpern, 1988; Shadish, 1992; Gallegos, 1994; and Georghiou, 1995).
- In this empirical metaevaluation, the authors used the parallel criteria suggested by Guba and Lincoln (1989: 236-243). Depending on the paradigmatic orientation of the evaluation study under investigation, the original or the revised version was applied. Thus, there were two different criteria sets available and the authors were in the position to select the appropriate one. This procedure made empirical analysis more sensitive to the approaches the evaluations followed. Evidently, this dualism is not able to catch the diversity of

epistemologies. On the other hand, it also restricts the comparability of the case studies elaborated with the different criteria sets. The dilemma between flexibility and sensitivity on the one hand, and comparability on the other, is a subject for further discussion.

The suggestions are not meant as a critique of the work done by the Joint Committee. They are intended as input to the ongoing discussion of context sensitive evaluation standard formation (compare Marklund, 1984; Raven, 1984; Nevo, 1984; Smith et al., 1993; Hendricks and Conner, 1995). As is probably the case other countries, Switzerland needs a set of criteria defining good practice in evaluation which takes its specific situation into account. *The Program Evaluation Standards* are definitely a helpful starting point for the Swiss discussion.

References

- Arbeitsgruppe 'Gesetzesevaluation' (AGEVAL) (1991) *Die Wirkungen staatlichen Handelns besser ermitteln: Probleme, Möglichkeiten, Vorschläge. Schlussbericht an das Eidgenössische Justiz- und Polizeidepartement. Oktober 1991*. Berne: EJPD/EDMZ.
- Balthasar, A. & Knöpfel, C. (1994) *Umweltpolitik und technische Entwicklung. Eine politikwissenschaftliche Evaluation am Beispiel der Heizungen*. Basle: Helbing & Lichtenhahn.
- Bierter, W. & Binder, H. (1993) *Unternehmerische Innovationsprozesse. Die Wirksamkeit staatlicher Förderungsmassnahmen - eine Evaluation*. Chur/Zurich: Rüegger.
- Bussmann, W. (1995) *Evaluationen staatlicher Massnahmen erfolgreich begleiten und nutzen. Ein Leitfaden*. Chur: Rüegger.
- Bussmann, W. (1996) Democracy and Evaluation's Contribution to Negotiation, Empowerment and Information: Some Findings From Swiss Democratic Experience. *Evaluation* 2(3): 307-319.
- Bussmann, W., Klöti, U. & Knoepfel, P. (1996) *Handbuch. Methoden der Evaluationsforschung*. Basle: Helbing & Lichtenhahn. (forthcoming)
- Forster, S. & Jeanrenaud, C. (1993) *La promotion économique neuchâteloise. Bilan et perspectives. Novembre 1993*. Neuchâtel: IRER, Université de Neuchâtel.
- Gallegos, A. (1994) Meta-evaluation of School Evaluation Models. *Studies in Educational Evaluation* 20: 41-54.
- Georgiou, L. (1995) Assessing the Framework Programmes. A Meta-evaluation. *Evaluation* 1(2): 171-188.
- Gerheuser, F. & Schmid, A. (1993) *Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes. Serie 'Schlussberichte' NFP 27*. Berne: NFP 27.
- Haering, B. et al. (1990) *Evaluation Kehricksackgebühr. Expertenbericht von Infrac. November 1990*. Berne: BUWAL.
- Hansen, J. (1994) The Joint Committee on Standards for Educational Evaluations. (1994). The Program Evaluation Standards: How to Assess Evaluations of Educational Programs, Thousand Oaks, CA: Sage. Book Review', *Educational and Psychological Measurement* 54(2): 550-567.
- Hendricks, M. & Conner, R. (1995) International Perspectives on the Guiding Principles, in W.R. Shadish, D.L. Newman, M.A. Scheirer and C. Wye, Christopher (eds.) *Guiding Principles*

for Evaluators. New Directions for Program Evaluation, No. 66, Summer 1995, pp. 77-90. San Francisco: Jossey-Bass Publishers.

Horber-Papazian, K. & Thèvoz, L. (1990) Switzerland: Moving towards Evaluation, in: R.C. Rist (ed.) *Program Evaluation and the Management of Government*, pp. 133-143. New Brunswick: Transaction Publishers.

Joint Committee on Standards for Educational Evaluation, ed. (1981) *Standards for Evaluations of Educational Programs, Projects, and Materials*. New York: McGraw-Hill Book Company.

Joint Committee on Standards for Educational Evaluation, ed. (1994) *The Program Evaluation Standards. How to Assess Evaluations of Educational Programs. 2nd Edition*. Thousand Oaks: Sage.

Klöti, U. (1987) Swiss Democracy. Exception or Model? Zurich. Unpublished manuscript.

Klöti, U (1988) Switzerland, in D.C. Rowat (ed.) *Public Administration in Developed Democracies. A Comparative Study*, pp. 189-203. New York: Marcel Dekker.

Knoepfel, P. et al. (1991) *Evaluation des BUWAL. Expertenbericht zur Evaluation der Luftreinhaltung, des ländlichen Gewässerschutzes und der UVP des Bundes*. Berne: EDMZ.

Linder, W (1994) *Swiss Democracy. Possible Solutions to Conflict in Multicultural Societies*. New York: St. Martin's Press.

Linder, W. et al. (1990) *Verbrauchsabhängige Heizkostenabrechnung, Förderung erneuerbarer Energien. Evaluation von Prozess und Wirkungen in drei Kantonen*. März 1990. Berne: FSP.

Longchamp, C. (1993) *Evaluierung der bundeseigenen EWR-Informationskampagne. Schlussbericht*. Berne: GfS.

Marklund, S. (1984) Applicability of Standards for Evaluations of Educational Programs, Projects, and Materials in an International Setting. *Evaluation and Program Planning* 7: 355-362.

Näf, A. et al. (1993) *Evaluation der Zürcher Gemeinschaftszentren (ZGZ). Vom Januar bis Mai 1993. Im Auftrag des Sozialamtes der Stadt Zürich*. Berne: Price Waterhouse.

Nevo, D. (1984) Applying the Evaluation Standards in a Different Social Context. *Evaluation and Program Planning* 7: 383-386.

Patton, M. (1994) The Program Evaluation Standards: How to Assess Evaluations of Educational Programs, by The Joint Committee on Standards for Educational Evaluations, Newbury Park, CA: Sage, (In Press). Book Review', *Evaluation Practice* 15(2): 193-199.

Raven, J. (1984) Some Limitations of the Standards. *Evaluation and Program Planning* 7: 363-370.

Schnyder, S. (1993) Über die Verteilung und Wirkung der Investitionshilfen im landwirtschaftlichen Hochbau des Kantons Zürich (zwischen 1970 und 1990). November 1993. Zurich. Unpublished manuscript.

Schulz, H, Christoph, M. and Jörg, H (1993) *Wohneigentumsförderung durch den Bund. Die Wirksamkeit des Wohnbau- und Eigentumsförderungsgesetzes (WEG). Schriftenreihe Wohnungswesen, Band 55*. Berne: BWO/EDMZ.

Schwandt, T. & Halpern, E. (1988) *Linking Auditing and Metaevaluation. Enhancing Quality in Applied Research*. Newbury Park: Sage.

Scriven, M. (1969) An Introduction to Meta-evaluation. *Educational Product Report* 2(5): 36-38.

Shadish, W. (1992) Theory-Driven Metaevaluation, in H.-T. Chen and P. H. Rossi (eds.) *Using Theory to Improve Program and Policy Evaluations*, pp. 29-45. New York: Greenwood Press.

Shadish, W., Newman, D., Scheirer, M., & Wye, C. eds. (1995) *Guiding Principles for Evaluators. New Directions for Program Evaluation, No. 66, Summer 1995*. San Francisco: Jossey-Bass Publishers.

Smith, M. & Glass, G. (1981) *Meta-Analysis of Social Research*. Beverly Hills: Sage.

Smith, N. (1981) Criticism and Meta-evaluation, in Nick L. Smith (ed.) *New Techniques for Evaluation*, pp. 266-274. Beverly Hills: Sage.

Smith, N., Chircop, S., & Mukherjee, P. (1993) Considerations on the Development of Culturally Relevant Standards, *Studies in Educational Evaluation*, 19(1): 3-13.

Sommer, H. & Suter, S. (1993) *Schneeräumung von Urner Passstrassen. Evaluation im Auftrag der Baudirektion des Kantons Uri, mitfinanziert durch NFP 27. Schlussbericht. September 1993*. Altdorf: Ecoplan.

Stanfield, J. & Smith, C. (1984) Management Consulting and Evaluation, *Evaluation and Program Planning* 7: 87-93.

Stufflebeam, D. (1978) Meta Evaluation: An Overview, *Evaluation & the Health Profession* 1: 17-43.

Suter, S. & Walter, F. (1993) *Gestaltungsmöglichkeiten und Auswirkungen verschiedener Submissionsordnungen. Evaluation und Bericht im Auftrag der Bau-, Verkehrs- und*

Energiedirektion des Kantons Bern, mitfinanziert durch NFP 27 und Eidg. Finanzverwaltung. Dezember 1993. Berne Ecoplan.

Thierstein, A. (1994) *Design Preis Schweiz. Eine Evaluation. September 1994. St. Gall: SIASR/HSG.*

Widmer, T. (1996a) *Meta-Evaluation: Kriterien zur Bewertung von Evaluationen. Berne: Haupt.*

Widmer, T. (1996b) *Fallstudien zur Meta-Evaluation, 7 Vols., Serie 'Schlussberichte' NFP 27. Berne: NFP 27.*

Widmer, T., Rothmayr, C., & Serdült, U. (1996) *Kurz und gut? Qualität und Effizienz von Kurzevaluationen. Zurich/Chur: Rüegger.*

Wuilloud, C. (1994) *Evaluation staatlicher Massnahmen. Effizienz der Massnahmen zur Abwehr von Naturgefahren im Kanton Wallis. Sitten. Unpublished manuscript.*